

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Frislin, Michael J.
Fingler, Douglas J.
Picarella, Dominic
Levman, Walter
- (ii) TITLE OF INVENTION: Mucosal Vascular Addressins and Uses Thereof
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
(B) STREET: Two Elm Street Drive
(C) CITY: Lexington
(D) STATE: Massachusetts
(E) COUNTRY: U.S.A.
(F) TEL: 617-451-4779
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patent In Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08,875,849
(B) FILING DATE: 11-FEB-1996
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08,523,004
(B) FILING DATE: 11-SEP-1995
- (viii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08,386,357
(B) FILING DATE: 10-FEB-1995
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Pross, David E.
(B) REGISTRATION NUMBER: 21,592
(C) REFERENCE, DOCKET NUMBER: LKS94-04A2
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 781-861-6240
(B) TELEFAX: 781-861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1624 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(12) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1218

(21) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAT TTC GGA CTC GCG CTC CTG CTG GCG GGG CTT CTG GGG CTC CTC 48
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
1 10 15

CTC GCG CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG 96
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu
20 25 30

CCG GTG GTG GAG GTG GCG TTG GGG GCG TCG CGC CAG CTC ACC TGC CGC 144
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg
35 40 45

CTG GCG TGC GAG GAG GCG GGG GCG TCG GTG CAG TGG CGG GGC CTC GAC 192
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp
50 55 60

ACC AGC CTG GAG GCG GAG CAG TCG GAG ACC GGC CGC AGC GTC CTC ACC 240
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr
65 70 75 80

GTG CGC AAC GCG TCG CTG TCG CTC GCG GCG ACC CGC GTG TGC GTG GCG 288
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly
85 90 95

TCC TGC GGG GGC CGC ACC TTC CAG CAC ACC GTG CAG CTC CTT GTG TAC 336
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr
100 105 110

GCC TTC CCG GAC CAG CTG ACC GTG TCC CCA GCA ACC CTC GTG CCG GGT 384
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly
115 120 125

GAC CCG GAG GCG GCG TGT ACC GCG CAC AAA GTC ACC CCC GTG GAC CCC 432
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro
130 135 140

AAC GCG CTC TCC TTC TTC CTG CTC GTC GCG GCG CAG GAG CTG GAG GCG 480
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly
145 150 155 160

GGG CAA GCG CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	628
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Glu Glu Glu Glu	
165 170 175	
GGG GAG GAG GAG GTG CTG TTC AGG GTG ACA GAG GCG TGG GCG CTG GCG GCG	676
Gly Asp Gln Asp Val Leu Phe Arg Val Thr Gln Ala Trp Arg Leu Phe	
180 185 190	
GGG CTG GGG ACC CCT GTC CCG CCC GCG CTC TAC TGC CAG GCG AGG ATG	674
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met	
195 200 205	
AGG CTG CCT GGC TTG CAG CTC AGC CAC CCG CAG GCG ATC CCC GTC CTG	672
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu	
210 215 220	
CAC AGC CCG ACC TCC CCG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	720
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
225 230 235 240	
CCC AAC ACC ACC TCC CCG GAG TCT CCC GAC ACC ACC TCC CCG GAG TCT	768
Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser	
245 250 255	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CAG GAG CCT	816
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro	
260 265 270	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	864
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
275 280 285	
CCC GAC AAG ACC TCC CCG GAG CTC GCG CCC CAG CAG GCG TCC ACA CAC	912
Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His	
290 295 300	
ACC CCC AAG ACC TCC CCG GAG CTC GCG CCC CAG CAG GCG TCC ACA CAC	960
Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser	
305 310 315 320	
CAG GCT GGG CCC ACC CAG GGA GAA GTG ATC CCA ACA GCG TCC TCC AAA	1008
Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys	
325 330 335	
CCT GCG GGT GAC CAG CTG CCG GCG GGT CTG TGG ACC AGC AGT GCG GTG	1056
Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val	
340 345 350	
CTG GGA CTG CTG CTC CTG GCG TTG CCC ACG TAT CAC CTC TGG AAA CCG	1104
Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg	
355 360 365	
TGC CGG CAC CTG GCT GAG GAC GAC ACC CAC CCA CCA GCT TCT CTG AGG	1152
Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg	
370 375 380	

mb
D
nt

B

ATT CTG CCC CAG GTG TCG GCG TGG GCT GGG TTA AGG GGG ACC GGC CAG 1200
 Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln
 121 129 136 140
 CTC GGG ATC AGC CCC TCC TGAGTGGCCA GCCTTTCCGC CTGTGAAAG 1248
 Val Gly Ile Ser Pro Ser
 125 135
 AAAATAGCTT GGACCCCTTC AAGTTGAGAA CTGGTCAGGG CAAACCTGCC TCCATTCTA 1388
 CTCAAAAGTCA TCCCTCTGCT CACAGAGATG GATGCATGTT CTGATTGCCCT CTTTGGAGAA 1408
 GCTCATCAGA AACTCAAAAG AAGGCCACTG TTTGTCTCAG CTACCCATGA CTTGAAGCC 1428
 CTCCCTGAGT GGTCGCCACC TTCTGGACG GAACCACGTA CTTTTTACAT ACATTGATTC 1488
 ATGTCTCAGG TCTCCCTAAA AATGCCAAG ACCAAGCTGT GCCTTGACCA CCTGCGGCT 1548
 CTCTCGTCAG GACCTCCTGA GCCTTTGGCA AATAAACCTC CTAAATGAT AAAAAAAAAA 1608
 AAAAAAAAAA AA/AAA 1624

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- A) LENGTH: 406 amino acids
- B) TYPE: amino acid
- C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu 15
 1 5 10
 Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu 30
 20 25 30
 Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg 45
 35 40 45
 Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp 60
 50 55 60
 Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr 80
 65 70 75 80
 Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly 95
 85 90 95
 Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr 110
 100 105 110
 Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly 125
 115 120 125

Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro
 130 135 140
 Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly
 145 150 155 160
 Ala Glu Ala Leu Gly Pro Glu Val Gln Glu Glu Gln Glu Glu Pro Gln
 165 170 175
 Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro
 180 185 190
 Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met
 195 200 205
 Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu
 210 215 220
 His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro
 225 230 235 240
 Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser
 245 250 255
 Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro
 260 265 270
 Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro
 275 280 285
 Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His
 290 295 300
 Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser
 305 310 315 320
 Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys
 325 330 335
 Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val
 340 345 350
 Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg
 355 360 365
 Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg
 370 375 380
 Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln
 385 390 395 400
 Val Gly Ile Ser Pro Ser
 405

126
 127
 128
 129
 130

131
 132
 133
 134
 135

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) FEATURES:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1146

(iv) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GAT TTC GGA CTG GCG CTG CTG GCG GGG CTT CTG GGG CTC CTC 48
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
1 5 10 15

CTC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG 96
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu
2 25 30

CCG GTG GTG GCG GTG GCG TTG GGC GCG TCG CGC CAG CTC ACC TGC CGC 144
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg
35 40 45

CTG GCG TGC GCG GAC CGC GCG GCG TCG CTG CAG TGG CGG GGC CTG GAC 192
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp
50 55 60

ACC AGC CTG GCG GCG GTG CAG TCG GAC ACC GCG CGC AGC GTC CTG ACC 240
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr
65 70 75 80

GTG CGC AAC GCG TCG CTG TCG GCG GCG GGG ACC GCG GTG TGC GTG GGC 288
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly
85 90 95

TCC TGG GGG GCG CGC ACC TTC CAG CAG ACC GTG CAG CTC CTT GTG TAC 336
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr
100 105 110

GCG TTC CTG GAC CAG CTG ACC TTC TCC CTA GCA GCG CTG GTG CCT GGT 384
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly
115 120 125

GAC CCG GAG GTG GCC TGT ACG GCG CAC AAA GTC ACG CCC GTG GAC CCC 432
Asp Phe Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro
130 135 140

AAC GCG CTC TCC TTC TTC CTG CTC CTC GGG GCG CAG GAA CTG GAG GGG 480
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly
145 150 155 160

GAG CAA GAG CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG GAG GAG
 Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Glu Glu
 185 190 195

GAG GAG GAG GAG GTG CTG TTC AGG GTG ACA GAG GAG GAG GAG GAG GAG GAG
 Gly Asp Glu Asp Val Leu Ile Arg Val Thr Gln Ala Trp Arg Leu Pro
 195 200 205

GAG CTG GGC ACC CCT CTC CCG CCG GCG CTC TAC TGG CAG GGT AGG ATG
 Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met
 195 200 205

AGG CTG CCT GGC TTG GAG CTC AGC CAC CCG CAG GCG ATC CCG CTC CTG
 Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu
 210 215 220

CAC AGC CCG ACC TCC CCG GAG CCT CCG GAC ACC ACC TCC CCG CAG TCT
 His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Ser
 225 230 235 240

CCG GAC ACC ACC TCC CCG GAG TCT CCG GAC ACC ACC TCC CAG GAG CCT
 Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro
 245 250 255

CCC GAC ACC ACC TCC CCG GAG CCT CCG GAC AAG ACC TCC CCG GAG CCC
 Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro
 260 265 270

GCC CCC CAG CAG GGC TCC ACA CAG ACC CCG AGG AGC CCA GGC TCC ACC
 Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr
 275 280 285

AGG ACT CGC CGC CCT GAG ATC TCC CAG GGT GGC CCG ACC CAG GGA GAA
 Arg Thr Arg Arg Pro Gln Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu
 290 295 300

GTG ATC CCA ACA GGT TCG TCC AAA CCT GCG GGT CAA CAG CTG CCG GCG
 Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala
 305 310 315 320

GGT CTG TGG ACC AGT AGT GCG GTG CTG GGA CTG CTG TCC CTG GCG TTG
 Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu
 325 330 335

CCC ACC TAT CAC CCG TGG AAA CCG TCC CCG CAG CTG CCT GAG GAG GAG
 Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp
 340 345 350

ACC CAG CCA CCA GGT TCT CTG AGC CTT CTG CCG CAG CTG TCG GCC TGG
 Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp
 355 360 365

GGT GGG TTA AGS GGG ACC GGC CAG GTC GGG ATC AGC CCG TCC
 Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser
 370 375 380

TGAGTGGCCA GCTTTTCC CTGTGAAAGC AAAATAGCTT GACCCCTT AAGTTGAGAA
 1206

CTGGTCAGGG CAAAGCTCC TCCATTCTA CTCAAAGTCA TCCCTCTGTT CACAGAGATG
 1266

GATGCATGTT CTGATTGCCT CTTTGGAGAA GTCATCAGA AACTCAAAAG AAGGCAATG 1336
 TTTCTTAA CTAGGATCA CTGAAGCCC CTGCTGACT GGTGCCCAGG TTCTTGAAG 1386
 GAAGCAGCTA CTTTTACAT ACATTGATTC ATGTCTCAGG TCTCCCTAAA AATGCTAAG 1446
 AGAAGCTGT GCTCTACCA CTTCTGCCC CTCTCTGAG GAGCTCTGA GCTTTPGGA 1506
 AATAAACCTC CTAAATGAA AAAAAAAAAA AAA 1559

(2) INFORMATION FOR SEQ ID NO:4:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 382 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(11) MOLECULE TYPE: protein

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu 1 5 10 15
 Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu 20 25 30
 Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg 35 40 45
 Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp 50 55 60
 Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr 65 70 75 80
 Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly 85 90 95
 Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr 100 105 110
 Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly 115 120 125
 Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro 130 135 140
 Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly 145 150 155 160
 Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Pro Gln 165 170 175
 Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro 180 185 190

Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met
195 200 205

Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu
210 215 220

His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Ser
225 230 235 240

Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro
245 250 255

Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro
260 265 270

Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr
275 280 285

Arg Thr Arg Arg Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu
290 295 300

Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala
305 310 315 320

Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu
325 330 335

Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp
340 345 350

Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp
355 360 365

Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser
370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1721 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 4..1638

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGC ATG GAT CGG GGC CTG ACC CTC CTG CTG GCG GGG CTT CTG GGG CTC
 Met Asp Arg Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu
 1 5 10 15

CTC CAG CCC GGC TGC GGC CAG TCC CTC CAG GTG AAG CCC CTC CAG GTG	36
Leu Gln Pro Gly Cys Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val	
20 35 50	
GAG CCC CCG GAG CCG GTG GTG GGC CTC GTC GTC GTC GTC CAG CAG	144
Gln Pro Pro Gln Pro Val Val Ala Val Ala Leu Gly Ala Ser Ala Ile	
35 40 45	
CTC ACC TGC CCG CTG GAC TCC CCG CAC CCC GGG GGC ACC GTG CAG TCG	148
Leu Thr Cys Arg Leu Asp Cys Ala Asp Arg Gly Ala Thr Val Gln Trp	
50 55 60	
CGG GGC CTC GAG ACC AGG CTG GGC GCG GTG CAG TCG GAC GCG GGC CCG	240
Arg Gly Leu Asp Thr Ser Leu Gly Ala Val Gln Ser Asp Ala Gly Arg	
65 70 75	
AGC GTC CTC ACC GTG CCG AAC GCC TCG CTC TCG CCG GCC GCG ACC CGT	288
Ser Val Leu Thr Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg	
80 85 90 95	
GTC TGC GTC GGC TCC TGC GGG GGC CCG ACC TTC CAG CAC ACC GTG CCG	336
Val Cys Val Gly Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Arg	
100 105 110	
CTC CTT GTG TAC GC : TTC CCG GAG CAG CTG ACC ATC TCC CCG GCA GCC	384
Leu Leu Val Tyr Ala Phe Pro Asp Gln Leu Thr Ile Ser Pro Ala Ala	
115 120 125	
CTG GTG CCT GGT GAG CCG GAG GTG GCG TGT ACG GCG CAC AAA GTC ACG	432
Leu Val Pro Gly Asp Pro Ser Val Ala Cys Thr Ala His Lys Val Thr	
130 135 140	
CCT GTG GAG CCG AAT GCG CTC TCC TTC TCC CTC CTG CTG CTG GGG GAC CAG	480
Pro Val Asp Pro Asn Ala Leu Ser Phe Ser Leu Leu Leu Gly Asp Gln	
145 150 155	
GAA CTG GAG GCG GCG CAG GGT CTG GAG CTG GAG GTG GAG GAG GAG GAG	528
Glu Leu Gln Gly Ala Gln Ala Leu Gly Pro Gly Val Gln Glu Glu Glu	
160 165 170 175	
GAG GAG CCC CAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	576
Glu Glu Pro Gln Glu Glu Glu Asp Val Leu Phe Arg Val Thr Glu Arg	
180 185 190	
FGG GAG CTC CCG ACC CTC GCA ACC CCT GTC CTC CCG GCG CTC TAC TGC	624
Trp Arg Leu Pro Thr Leu Ala Thr Pro Val Leu Pro Ala Leu Tyr Cys	
195 200 205	
CAG GCG ACC ATG AGG CTG CCT GCG TTC GAG CTC ACC CAG CTC CAG GCC	672
Gln Ala Thr Met Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala	
210 215 220	
ATC CCG GTG CTG CAG GCG ACC TTC GGG GAG CCG CCG GAG ACC ACC	720
Ile Pro Val Leu His Gly Pro Thr Ser Arg Glu Pro Pro Asp Thr Thr	
225 230 235	
TCC CCG GAA CCG CCG GCG GCG ACC TTC CCG GAG ACC ACC CCG CAG CAG	768
Ser Pro Glu Pro Arg Ala Ala Thr Ser Pro Glu Thr Thr Pro Gln Gln	
240 245 250 255	

Mc
D1

GGC TCC ACA CAC AGC CCC AGG AGC CCG CCC TCT ACC AGC AAT TGC CAG 816
 Gly Ser Thr His Ser Pro Arg Ser Pro Gly Ser Thr Ala Thr Cys Arg
 260 265 270
 CTT GAG ATG TCC CAG GCT GCG CCC AGG CAG GGA GAA GIG ATT CTA ATA 864
 Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr
 275 280 285
 GGC TCC TCC AAA CCT ACC GGT GAC CAG CTG CCC GCG GCT CTG TCG ACC 912
 Gly Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr
 290 295 300
 AGC AGT GCG GTG CTG CGA CTG CTG CTC CTG GCT TTG CCC ACC TAC CAC 960
 Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His
 305 310 315
 CTC TGG AAA CGT TGC CGG CAC CTG GCT GAG GAC GGC CCC CAC CCA CCA 1008
 Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro
 320 325 330 335
 GCT TCT CTG AGT AGC CAG CCC TCC CCC CTG TGAAGGGAAA ATAGGTTGGA 1058
 Ala Ser Leu Ser Ser Gln Pro Phe Pro Leu
 340 345
 CCCCTTCAAG CTGAGAGCTG GTGCGGGGAAA ACCCTGCCTCC CATTCTATTC AAAGTCATCG 1118
 CTCTGGTCAC AGAGAGGAC GCACATTCTG ATTGCCTCCT TTGGAAAGGC TCATCAGAAA 1178
 CTCAAAAGAA GTTATCTTT TGTTCGCTT ACCCGTGAAC TGGAAGCCCC CGCCCCGCTC 1238
 GAGTGACCCC TGAATTCTTG GAGGAAACCA ACGTACTTCT TACATATATT GATTTCATGTG 1298
 TCATATCTCC CTAATAATGCG TAAAGCAGC TGTGCTCCGA CCACCTTGGG CCCCTGCCAT 1358
 CAGGAACCTCC TGAGGCTTG GAAATAAAG CTCTAAAG GATAGAACT GAAACTTGTG 1418
 GCGGGGCGCG GTGGCTAAG CCGTAATCC CAGCACTTTG GGAGGCGAG GTGGTGGAT 1478
 CACGAGGTCA GGAGATGAG ACCATCTCTT CTAAACCGTG AAACCCCGT TCTACTAAAA 1538
 AAATACAAAA ATTAGCGGG AGCGGTGCG GGCGCTGTA GTCCAGCTA CTGGGAGGC 1598
 TGAAGCAGGA GAATGGCTG AACCGGAGG GCGGAGCTTG CAGTGAGCTG AGATCCGSCC 1658
 ACTGCACTCC AGCTGCGGG ACAGAGCGAG ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA 1718
 AAA 1721

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 345 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(41) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asp Arg Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
1 5 10 15
Gln Pro Gly Cys Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Gln
20 25 30
Pro Pro Gln Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu
35 40 45
Thr Cys Arg Leu Asp Cys Ala Asp Arg Gly Ala Thr Val Gln Trp Arg
50 55 60
Gly Leu Asp Thr Ser Leu Gly Ala Val Gln Ser Asp Ala Gly Arg Ser
65 70 75 80
Val Leu Thr Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val
85 90 95
Cys Val Gly Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Arg Leu
100 105 110
Leu Val Tyr Ala Phe Pro Asp Gln Leu Thr Ile Ser Pro Ala Ala Leu
115 120 125
Val Pro Gly Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro
130 135 140
Val Asp Pro Asn Ala Leu Ser Phe Ser Leu Leu Leu Gly Asp Gln Glu
145 150 155 160
Leu Glu Gly Ala Gln Ala Leu Gly Pro Glu Val Glu Glu Glu Glu Glu
165 170 175
Glu Pro Gln Glu Glu Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp
180 185 190
Arg Leu Pro Thr Leu Ala Thr Pro Val Leu Pro Ala Leu Tyr Cys Gln
195 200 205
Ala Thr Met Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile
210 215 220
Pro Val Leu His Gly Pro Thr Ser Arg Glu Pro Pro Asp Thr Thr Ser
225 230 235 240
Pro Glu Pro Arg Ala Ala Thr Ser Pro Glu Thr Thr Pro Gln Gln Gly
245 250 255
Ser Thr His Ser Pro Arg Ser Pro Gly Ser Thr Arg Thr Cys Arg Pro
260 265 270
Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly
275 280 285
Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser
290 295 300

Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu
305 310 315 320

Trp Lys Arg Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro Ala
325 330 335

Ser Leu Ser Ser Glu Pro Phe Pro Leu
340 345

(2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCTACTGCC AGGGAAG

18

(2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCCTGGGA3 ATCTAAGG

19

(2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCACGATGA AGCTGCTTG

20

(2) INFORMATION FOR SEQ ID NO:10:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly Leu Asp Thr Ser Leu
1 5

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ix) FEATURE:
A: NAME/KEY: Region
B: LOCATION: 2
D: OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Ile or Leu"

(ix) FEATURE:
A: NAME/KEY: Region
B: LOCATION: 3
D: OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Asp or Glu"

(ix) FEATURE:
A: NAME/KEY: Region
B: LOCATION: 4
D: OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Thr or Ser"

(ix) FEATURE:
A: NAME/KEY: Region
B: LOCATION: 5
D: OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Pro or Ser"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gly Xaa Xaa Xaa Xaa Leu
1 5

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

His Ile Asp Ser Pro Leu
1 5

(ii) INFORMATION FOR SEQ ID NO:17:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Leu Asp Thr Ser Leu
1 5